Traitement du Signal sur Graphe : Interprétation en termes de Filtre de l'Apprentissage Semi-Supervisé sur Graphe

Benjamin Girault, Eric Fleury, Paulo Gonçalves

Traitement du Signal sur Graphe : Interprétation en termes de Filtre de l'Apprentissage Semi-Supervisé sur Graphe

Benjamin Girault[1,3], Éric Fleury[1,2,3], Paulo Gonçalves[2,3]

[1]*École Normale Supérieure de Lyon*
[2]*Inria*
[3]*LIP (UMR CNRS – ENS Lyon – UCB Lyon 1 – Inria 5668), 46 allée d'Italie, 69364 Lyon Cedex 7, France*

Nous montrons comment les outils de traitement du signal sur graphe peuvent dégager des notions de fréquences sur les graphes pour étudier des données portées par les nœuds d'un graphe. Prenant l'exemple de l'apprentissage semi-supervisé, nous montrons alors qu'il peut s'interpréter comme le filtre d'un signal sur graphe.

Keywords: apprentissage semi-supervisé, network science, traitement du signal sur graphe

1 Introduction

L'avènement des moyens techniques de collecte de données tels que les réseaux de capteurs a permis de rassembler rapidement des données précises et dynamiques sur des réseaux divers. On citera par exemple la charge d'ordinateurs en réseau, la température relevée par un réseau de capteurs, le nombre de messages publiés par les utilisateurs d'un réseau social, la densité de protéines présentes dans un milieu pour un réseau d'interaction protéine-protéine, ou simplement les couleurs des pixels d'une image.

Il s'avère que comprendre ces données nécessite de tenir compte des relations entre ses entités du fait des corrélations qu'elles induisent. C'est ainsi que la température relevée par un capteur varie peu entre deux capteurs proches dans l'espace. Malheureusement, ces structures de réseau ne sont pas toujours représentables par des espaces euclidiens pour lesquels les données peuvent être étudiées par les outils classiques du traitement du signal. Dès lors, il est naturel de vouloir généraliser ces outils sur de telles structures non-euclidiennes afin de pouvoir décomposer un signal en composantes « fréquentielles » isolant les variations globales des variations locales, *i.e.* les « basses fréquences » des « hautes fréquences ».

Les auteurs de [SNF$^+$13] font un état de l'art détaillé du domaine que nous ne prétendons pas résumer ici. En revanche, deux travaux récents méritent d'être cités. Tout d'abord l'approche ondelettes sur graphe développée dans [HVG11] permet de décomposer un signal en composantes centrées sur un nœud donné et à une échelle donnée, afin par exemple de modéliser des phénomènes de diffusion (voir [GGF13]). La deuxième approche est celle développée dans [SM13], que nous utiliserons ici, définissant une notion de translation sur graphe, l'équivalent d'une translation en temps, et une décomposition fréquentielle associée. Il est à noter que ces approches, ainsi que celles présentées dans [SNF$^+$13], se concentrent sur des graphes et des signaux statiques et qu'étant donnée une application la question de l'approche la plus adaptée reste ouverte.

Pour illustrer l'intérêt de l'approche traitement du signal, nous proposons l'étude d'un outil d'analyse de graphe particulier : l'apprentissage semi-supervisé sur graphe introduit dans le cadre de la classification. Cet outil part d'une connaissance et la transforme pour obtenir une classification. Nous montrons ici que cet a priori peut être interprété comme un signal et la transformation comme un filtre.

Le reste de ce document présente en section 2 un cadre méthodologique de l'apprentissage semi-supervisé sur graphe, puis du traitement du signal sur graphe en section 3. Enfin, la section 4 présente l'interprétation de l'apprentissage semi-supervisé comme un filtre de signal sur graphe et un exemple simple d'application.

2 Apprentissage Semi-Supervisé sur Graphe

Nous allons montrer que les méthodes classiques d'apprentissage semi-supervisé sur graphe (noté *ASSG* par la suite) peuvent s'interpréter comme des filtres de signaux sur graphe dans la théorie récemment développée du traitement du signal sur graphe. Pour cela, nous étudions la méthode d'ASSG développée dans [AGMS12] qui présente l'avantage de faire le lien entre trois méthodes classiques que sont la méthode standard, la méthode normalisée, et la méthode PageRank, via un unique paramètre d'unification σ.

Ces méthodes d'ASSG ont été introduites pour classifier les nœuds d'un graphe. Plus particulièrement, étant donnés un graphe pondéré (non dirigé) de N nœuds, un ensemble de nœuds pré-classés dans K classes, la classification consiste à assigner exactement une classe à chaque nœud en tenant compte du poids des arêtes reliant ces nœuds.

Pour cela, la méthode présentée dans [AGMS12] considère la matrice Y de N lignes et K colonnes telle que $Y_{n,k} = 1$ lorsque le nœud n est pré-classé dans la classe k, et $Y_{n,k} = 0$ sinon. Étant donnés D la matrice diagonale des degrés des nœuds, A la matrice des poids des arêtes, et $\mathscr{L} = D - A$ le laplacien standard du graphe, il s'agit alors de résoudre le problème d'optimisation suivant :

$$\min_F \Big\{ \sum_{k=1}^{K} \underbrace{\|D^{\sigma-\frac{1}{2}}(F_{.k} - Y_{.k})\|_2^2}_{\text{ajustement}} + \alpha \underbrace{\|D^{\sigma-\frac{1}{2}} F_{.k}\|_{\mathscr{L}}^2}_{\text{pénalité}} \Big\}, \tag{1}$$

où $\|X\|_{\mathscr{L}} = X^T \mathscr{L} X$ est la semi-norme Laplacienne. Ce problème d'optimisation calcule un compromis entre un terme d'ajustement aux données et un terme de pénalité forçant la régularité de la solution sur le graphe. Ce compromis est dirigé par le paramètre de régularisation α.

Cette méthode présente l'avantage d'avoir une solution unique sous forme analytique :

$$\begin{array}{rrcl} & F & = & C_{\sigma,\alpha} Y \\ \text{tel que} & {C_{\sigma,\alpha}}^{-1} & = & (1+\alpha) I_N - \alpha D^{-\sigma} A D^{\sigma-1}. \end{array} \tag{2}$$

Pour un problème de classification, chaque nœud n est alors classé dans la classe k de score maximal $F_{n,k}$.

3 Traitement du Signal sur Graphe

Nous présentons maintenant brièvement la théorie du traitement du signal sur graphe développée dans [SM13]. Cette théorie transpose la notion de filtre linéaire invariant par translation (noté *LIT* par la suite), classe la plus commune de filtres. Nous savons que pour des signaux temporels, la sortie de ces filtres peut s'exprimer comme la convolution du signal par la réponse impulsionnelle du filtre (*i.e.* la réponse du filtre à un signal Dirac). Caractériser un tel filtre est donc très simple. Dans la théorie des filtres LIT, une notion centrale est la notion de *translation*. En effet, un filtre invariant par translation est un opérateur commutant avec l'opérateur de translation. Un filtre est alors LIT si et seulement si il s'exprime comme un polynôme en cet opérateur de translation.

[SM13] propose comme opérateur de translation sur graphe la multiplication de la matrice des poids des arêtes par le signal. Cette opération propage une mesure localisée sur un nœud à ses voisins immédiats. Dès lors, un filtre linéaire H sur le graphe s'exprimant comme une matrice carrée de taille N, H est invariant par translation si et seulement si A et H commutent. En particulier, c'est le cas si H peut s'exprimer comme un polynôme en la matrice A. En parlera de filtre linéaire invariant par translation de graphe (noté filtre *LITG* par la suite).

De plus, la théorie des filtres LIT montre que la convolution du signal par la réponse impulsionnelle du filtre est équivalente à la multiplication de la transformée de Fourier du signal par la réponse fréquentielle du filtre, *i.e.* l'atténuation du signal par le filtre en fonction des fréquences. Partant de ce constat, les auteurs de [SM13] utilisent la décomposition spectrale de la matrice de poids $A = U \Lambda U^{-1}$ † et définissent la transformée de Fourier d'un signal X comme $\widehat{X} = U^{-1} X$, pour définir une propriété analogue.

La réponse fréquentielle d'un filtre LITG devient alors un polynôme en Λ. En outre, la multiplication du signal par la matrice $h(A)$ devient la multiplication de la transformée de Fourier du signal par la matrice

†. Si la matrice n'est pas diagonalisable, la Jordanisation est utilisée. Voir [SM13] pour plus de détails.

diagonale $h(\Lambda)$. On pourra reconnaître ici la transposition du théorème de convolution aux signaux sur graphe. On peut observer par exemple que si $h = 1$, alors $h(A) = I_N = h(\Lambda)$, donc le filtre n'a aucune influence et toutes les fréquences sont conservées. Si par contre, h est une fonction de λ tendant vers 0, alors plus la valeur propre λ_i est grande, et plus la composante fréquentielle du signal associée à λ_i est atténuée. Ce filtre a donc un effet de lissage passe-bas, les fréquences basses étant conservées et les fréquences hautes coupées.

Nous avons donc, pour un graphe donné, une notion de translation, une notion de décomposition spectrale d'un signal, et une notion d'équivalence que l'on exprime un filtre dans le domaine direct ou dans le domaine spectral.

4 Interprétation comme Filtre de l'Apprentissage Semi-Supervisé

Dans cette section, nous appliquons les méthodes présentées dans les sections 2 et 3 pour montrer que les méthodes d'ASSG s'interprètent comme des filtres de signaux sur graphe. Pour cela, nous utilisons le cas simple du cycle de taille N, cas d'école présenté en détail dans [GGFS14]. Ce cas simple présente l'intérêt d'être tel que $D = I_N$, si bien que le paramètre d'unification σ n'a pas d'influence. Ceci permet donc l'étude plus précise du paramètre α, qui s'avère caractéristique de la bande passante du filtre équivalent.

De plus, ce cas permet d'écrire la matrice $C_{\sigma,\alpha}$ comme un polynôme en A tel que pour toute valeur propre λ_i on a $h(\lambda_i) = ((1+\alpha) - \alpha\lambda_i)^{-1}$. Dès lors, en utilisant (2), ASSG est un filtre de signal sur graphe passe-bas, de bande passante d'autant plus étroite que α est grand (voir [GGFS14] pour les détails). Nous reconnaissons également dans l'expression de h l'expression d'un filtre de Wiener, qui est ici un estimateur d'un signal à partir d'une observation avec un bruit blanc de puissance spectrale $1/\alpha$ qui minimise l'erreur quadratique moyenne.

En outre, la matrice $C_{\sigma,\alpha}$ étant multipliée par la matrice Y définie en section 2 pour obtenir la matrice F, nous pouvons interpréter les matrices Y et F comme les signaux d'entrée et de sortie du filtre, chaque nœud étant le support d'un vecteur de dimension K. Considérons maintenant un signal X quelconque, possiblement bruité, de dimension 1 (a priori sur une classe). Ceci revient à placer une valeur sur chaque nœud du graphe. La matrice $C_{\sigma,\alpha}$ de l'ASSG vue comme un filtre passe-bas sur graphe peut alors être appliquée à ce signal pour obtenir un signal de sortie F.

La Figure 1 présente les résultats de cette approche, à la fois du point de vue de l'ASSG (ligne supérieure) que du point de vue du traitement du signal sur graphe (ligne inférieure). On peut en particulier observer l'action du paramètre de régularisation α sur le gabarit du filtre : plus α est petit et plus le filtre est *passe-tout*. Ceci s'explique grâce à (1) où pour α proche de 0, le terme de pénalité n'a aucune influence, et F sera proche de Y. En revanche pour α élevé, le filtre ne laissera passer que la *composante continue* du signal : le terme de pénalité est alors dominant, et la solution la plus lisse sur le graphe est sélectionnée.

5 Conclusion

On a pu voir qu'un outil classique tel que l'apprentissage semi-supervisé sur graphe peut se réinterpréter dans le cadre du traitement du signal sur graphe comme le filtrage d'un signal défini par la méthode, à savoir l'a priori sur les nœuds. Le traitement du signal sur graphe est alors un outil puissant au champ d'application extrêmement large. De plus, la notion même de fréquence sur un graphe permet de reconnaître diverses notion classiques sur les signaux sur graphes telles que le bruit, les discontinuités ou les tendances. La question de l'extension des résultats montrés dans le cas simple du cycle aux graphes plus compliqués reste maintenant à étudier.

La notion de translation sur graphe est au centre de communications récentes autour du traitement du signal sur graphe. On a vu [SM13] et on pourra citer [SRV13] pour une définition différente. Dès lors, avoir à disposition cet opérateur ouvre le champ théorique de la stationnarité de signaux sur graphe, celle-ci se définissant dans le cas classique par une invariance par translation. On attend alors d'une telle notion la compréhension de mesures stables sur un graphe. Si nous prenons l'exemple de la charge de nœuds d'un réseau informatique déjà évoqué en introduction, on attendra une stationnarité dépendante de l'hétérogénéité de la topologie. Nous travaillons actuellement sur une notion de stationnarité prenant en compte la topologie du réseau.

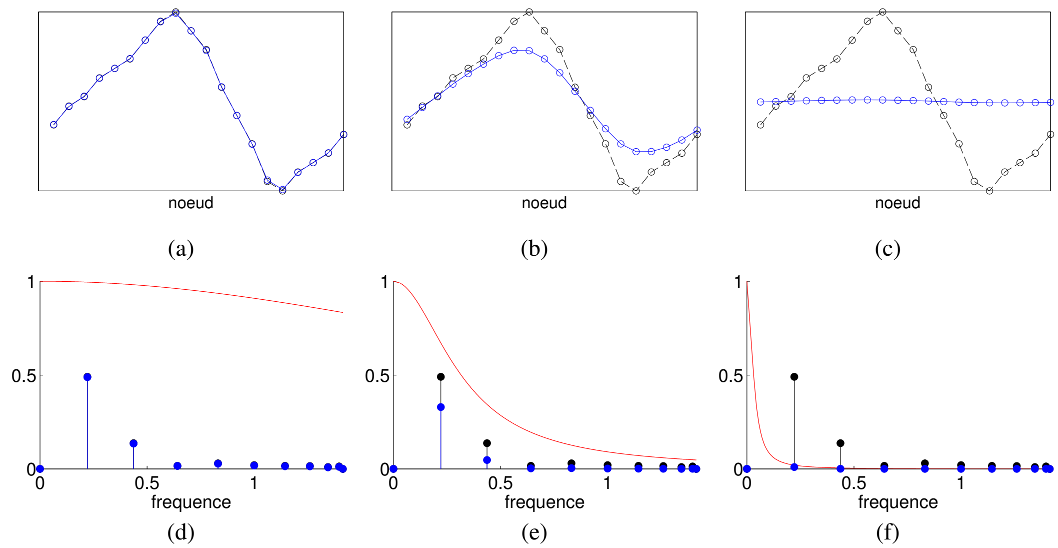

FIGURE 1: Les Figures (a)-(c) montrent les valeurs d'un signal bruité (pointillés noir) porté par un graphe cyclique, ainsi que le signal filtré (en bleu) par une méthode d'apprentissage semi-supervisé. Les trois colonnes correspondent à des paramètres de régularisation du plus faible au plus élevé ($\alpha = 0.1$, $\alpha = 10$ et $\alpha = 1000$) avec en haut la représentation directe, et en bas la représentation spectrale. Enfin, les Figures (d)-(f) montrent le gabarit du filtre (en rouge), *i.e.* l'atténuation du filtre en fonction de la fréquence, avec les énergies des composantes fréquentielles du signal bruité (noir) et du signal filtré (bleu).

Enfin, les réseaux sont rarement statiques, en particulier les réseaux sociaux. Dès lors, il serait intéressant d'obtenir des méthodes de traitement du signal appliqué à des signaux dynamiques sur de telles structures dynamiques. Quelle notion de fréquence peut on développer pour une analyse harmonique pertinente ?

Références

[AGMS12] K. AVRACHENKOV, P. GONÇALVES, A. MISHENIN et M. SOKOL : Generalized Optimization Framework for Graph-based Semi-supervised Learning. *In SDM*, pages 966–974. SIAM / Omnipress, 2012.

[GGF13] B. GIRAULT, P. GONÇALVES et É. FLEURY : Graphe de contacts et ondelettes : étude d'une diffusion bactérienne. *In Proceedings of Gretsi 2013*, septembre 2013.

[GGFS14] B. GIRAULT, P. GONÇALVES, E. FLEURY et A. SINGH MOR : Semi-Supervised Learning for Graph to Signal Mapping : a Graph Signal Wiener Filter Interpretation. *In Proc. of IEEE Int. Conf. on Acoustics, Speech, and Sig. Proc. (ICASSP)*, Florence, Italie, mai 2014. À paraître.

[HVG11] D. K. HAMMOND, P. VANDERGHEYNST et R. GRIBONVAL : Wavelets on graphs via spectral graph theory. *Applied and Computational Harmonic Analysis*, 30(2):129–150, 2011.

[SM13] A. SANDRYHAILA et J. M. F. MOURA : Discrete Signal Processing on Graphs. *IEEE Transactions on Signal Processing*, 61(7):1644–1656, 2013.

[SNF+13] D. I. SHUMAN, S. K. NARANG, P. FROSSARD, A. ORTEGA et P. VANDERGHEYNST : The Emerging Field of Signal Processing on Graphs : Extending High-Dimensional Data Analysis to Networks and Other Irregular Domains. *IEEE Signal Process. Mag.*, 30(3):83–98, 2013.

[SRV13] D. I. SHUMAN, B. RICAUD et P. VANDERGHEYNST : Vertex-Frequency Analysis on Graphs. *submitted to Applied and Computational Harmonic Analysis*, July 2013.